

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2005, 12:51:51 ; Search time 60.8 seconds
(without alignments)
58.956 Million cell updates/sec

Title: US-10-080-100-46
Perfect score: 33
Sequence: 1 Q9ANLRA 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 03.*
1: uniprot_sprot.*
2: uniprot_crembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	393	Q7NM92	Q7nm92 gloeobacter
2	33	100.0	599	Q9HCE8	Q9hce8 homo sapien
3	30	90.9	270	1 TRPA_BUCDN	Q8429 buchnera ap
4	30	90.9	270	1 TRPA_BUCMH	Q9r33 buchnera ap
5	30	90.9	271	2 Q6CH54	Q6ch54 yarrowia li
6	30	90.9	274	2 Q9S3U2	Q9s3u2 zymomonas m
7	30	90.9	656	2 Q89G6	Q89g6 pseudomonas
8	30	90.9	885	2 Q8BJ62	Q8bj62 debaryomyce
9	30	90.9	1018	2 Q7F1B4	Q7f1b4 anopheles g
10	29	87.9	158	2 Q647G8	Q647g8 arachis hyp
11	29	87.9	160	2 Q9SQH1	Q9sqh1 arachis hyp
12	29	87.9	245	2 Q8PQV7	Q8pqv7 xanthomonas
13	29	87.9	251	2 Q631T0	Q631t0 burkholderi
14	29	87.9	269	1 TRPA_BUCBP	Q9457 buchnera ap
15	29	87.9	280	2 Q7D767	Q7d767 vibrio vuln
16	29	87.9	503	2 Q6ZXCO	Q6zxc0 trypanosoma
17	29	87.9	573	2 Q81CT8	Q81ct8 bacillus ce
18	29	87.9	631	2 Q7NFP2	Q7nfp2 gloeobacter
19	29	87.9	748	2 Q9S8S9	Q9s8s9 bradyrhizob
20	29	87.9	804	2 Q7XL90	Q7xl90 oryza sativ
21	29	87.9	823	2 Q7X8R1	Q7x8r1 oryza sativ
22	29	87.9	877	2 Q7XKH6	Q7xkh6 oryza sativ
23	29	87.9	916	2 Q8L4K9	Q8l4k9 oryza sativ
24	29	87.9	1822	2 Q8VRG6	Q8vrg6 chloroflexu
25	28	84.8	97	2 Q857L9	Q857l9 mycobacteri
26	28	84.8	149	2 Q7ZZU3	Q7zzu3 oreochromis
27	28	84.8	175	2 Q9NZB5	Q9nzb5 homo sapien
28	28	84.8	192	2 Q8LKH3	Q8lkh3 brassica ol
29	28	84.8	217	2 Q856W9	Q856w9 mycobacteri
30	28	84.8	268	1 TRPA_SALTI	Q827e0 salmonella
31	28	84.8	268	1 TRPA_SALTY	P00929 salmonella

ALIGNMENTS

RESULT 1

ID	Q7NM92	Q7NM92	PRELIMINARY;	PRT;	393 AA.
AC	Q7NM92;				
DT	01-MAR-2004 (TrEMBLrel. 26, Created)				
DT	01-MAR-2004 (TrEMBLrel. 26, Last sequence update)				
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)				
DE	G110875 protein.				
GN	OrderedLocusNames=g110875;				
OS	Gloeobacter violaceus.				
OC	Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.				
OX	NCBI_TaxID=33072;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=PCC 7421;				
RX	MEDLINE=22977040; PubMed=14621292;				
RA	Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T., Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpo S., Takeuchi C., Yamada M., Tabata S.;				
RT	"Complete genome structure of Gloeobacter violaceus PCC 7421, a cyanobacterium that lacks thylakoids.";				
RL	DNA RES. 10:137-145(2003).				
DR	EMBL; AP006571; BAC88816.1; -				
DR	GO; GO:0016020; C:membrane; IEA.				
DR	GO; GO:0008565; F:protein transporter activity; IEA.				
DR	GO; GO:0009306; P:protein secretion; IEA.				
DR	InterPro; IPR011055; Dup.hybrid_motif.				
DR	InterPro; IPR006143; HlyD.				
DR	InterPro; IPR011053; Hybrid_motif.				
DR	Pfam; PF00529; HlyD; 1.				
DR	TIGRFAMs; TIGR01730; RND_mfp; 1.				
KW	Complete proteome.				
SK	SEQUENCE 393 AA; 42485 MW; 28EC63AC18B981CE CRC64;				

Query Match 100.0%; Score 33; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1 Q9ANLRA 7
Db	179 Q9ANLRA 185

RESULT 2

ID	Q9HCE8	PRELIMINARY;	PRT;	599 AA.
AC	Q9HCE8;			
DT	01-MAR-2001 (TrEMBLrel. 16, Created)			
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	KIAA1624 protein (Fragment).			
GN	Name=KIAA1624;			
OS	Homo sapiens (Human).			

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.";
RL DNA Res. 7:273-281(2000).
DR EMBL; AB046844; BAB13450.1; -.
DR GenBank; HGNC:17830; GPR107.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR InterPro; IPR009637; Lung_7TM_recept.
DR Pfam; PF06814; Lung_7-TM_R; 1.
FT NON_TER 1
SQ SEQUENCE 599 AA; 66906 MW; DC2AD69FAE078500 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 599;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QANLRA 7
Db 443 QANLRA 449

RESULT 3
TRPA_BUCDN STANDARD; PRT; 270 AA.
AC O68429;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39; Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Tryptophan synthase alpha chain (EC 4.2.1.20).
GN Name=trpA;
OS Buchnera aphidicola (subsp. Diuraphis noxia).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98290710; PubMed=9625791;
RA Baumann L., Baumann P., Moran N.A.;
RT "The endosymbiont (Buchnera) of the aphid Diuraphis noxia contains all
RT the genes of the tryptophan biosynthetic pathway.";
RL Curr. Microbiol. 37:58-59(1998).
CC -!- FUNCTION: The alpha subunit is responsible for the aldol cleavage
CC of indoleglycerol phosphate to indole and glyceraldehyde 3-
CC phosphate.
CC -!- CATALYTIC ACTIVITY: L-serine + 1-(indol-3-yl)glycerol 3-phosphate
CC = L-tryptophan + glyceraldehyde 3-phosphate.
CC -!- PATHWAY: Tryptophan biosynthesis; fifth (last) step.
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -!- SIMILARITY: Belongs to the trpA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF038565; AAC27736.1; -.
CC HSSP; P00929; 2WSY.
CC HAMAP; MF_00131; -.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF038565; AAC27736.1; -.
CC HSSP; P00929; 2WSY.
CC HAMAP; MF_00131; -.
CC
CC InterPro; IPR003009; FMN enzyme.
CC InterPro; IPR011060; RibP_bind_barrel.
CC Pfam; PF00290; Trp_syntA; 1.
DR
```

```
DR ProDom; PD001535; Trp_synthaseA; 1.
DR TIGRFAMs; TIGR00262; trpA; 1.
DR PROSITE; PS00167; TRP_SYNTHASE_ALPHA; 1.
KW Lyase; tryptophan biosynthesis.
FT ACT SITE 49 Proton acceptor (By similarity).
FT ACT SITE 60 Proton acceptor (By similarity).
SQ SEQUENCE 270 AA; 30442 MW; D8D3E73C4CD81CCD CRC64;

Query Match 90.9%; Score 30; DB 1; Length 270;
Best Local Similarity 85.7%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QANLRA 7
Db 65 QANLRA 71

RESULT 4
TRPA_BUCMH STANDARD; PRT; 270 AA.
AC Q8Q33;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41; Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Tryptophan synthase alpha chain (EC 4.2.1.20).
GN Name=trpA;
OS Buchnera aphidicola (subsp. Melaphis rhois).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20022990; PubMed=10555290;
RA Clark M.A., Moran N.A., Baumann P.;
RT "Sequence evolution in bacterial endosymbionts having extreme base
RT compositions.";
RL Mol. Biol. Evol. 16:1586-1598(1999).
CC -!- FUNCTION: The alpha subunit is responsible for the aldol cleavage
CC of indoleglycerol phosphate to indole and glyceraldehyde 3-
CC phosphate.
CC -!- CATALYTIC ACTIVITY: L-serine + 1-(indol-3-yl)glycerol 3-phosphate
CC = L-tryptophan + glyceraldehyde 3-phosphate.
CC -!- PATHWAY: Tryptophan biosynthesis; fifth (last) step.
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -!- SIMILARITY: Belongs to the trpA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF132318; AAF14255.1; -.
CC HSSP; P00929; 1K7E.
CC HAMAP; MF_00131; -.
CC
CC InterPro; IPR003009; FMN enzyme.
CC InterPro; IPR011060; RibP_bind_barrel.
CC InterPro; IPR002028; Trp_synthaseA.
CC Pfam; PF00290; Trp_syntA; 1.
DR ProDom; PD001535; Trp_synthaseA; 1.
DR TIGRFAMs; TIGR00262; trpA; 1.
DR PROSITE; PS00167; TRP_SYNTHASE_ALPHA; 1.
KW Lyase; tryptophan biosynthesis.
FT ACT SITE 49 Proton acceptor (By similarity).
FT ACT SITE 60 Proton acceptor (By similarity).
SQ SEQUENCE 270 AA; 30398 MW; 9FBECE911836C3EC CRC64;

Query Match 90.9%; Score 30; DB 1; Length 270;
Best Local Similarity 85.7%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy 1 QRANLRA 7
Db 65 QRANLRA 71

RESULT 5
Q6CH54 PRELIMINARY; PRT; 271 AA.
ID Q6CH54
AC Q6CH54;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Similarity.
GN ORFNames=YAL10A12221g;
OS Yarrowia lipolytica CLIB99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talia E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisarane A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44 (2004).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG Genoscope;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382127; CAG83937.1; -.
SQ SEQUENCE 271 AA; 29679 MW; B116E396233B4D94 CRC64;

Query Match 90.9%; Score 30; DB 2; Length 271;
Best Local Similarity 85.7%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLRA 7
Db 123 QRANLRS 129

RESULT 6
Q9S3U2 PRELIMINARY; PRT; 274 AA.
ID Q9S3U2
AC Q9S3U2;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Tryptophan synthetase alpha subunit.
GN Name=trpA;
OS Zymomonas mobilis (subsp. pomaceae).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Zymomonas.
OX NCBI_TaxID=120044;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29192;
RG Eddy C.K., Ingram L.O.;

```

```

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: The alpha subunit is responsible for the aldol cleavage
CC of indoleglycerol phosphate to indole and glyceraldehyde 3-
CC phosphate (By similarity).
CC -1- CATALYTIC ACTIVITY: L-serine + 1-(indol-3-yl)glycerol 3-phosphate
CC = L-tryptophan + glyceraldehyde 3-phosphate.
CC -1- PATHWAY: Tryptophan biosynthesis; fifth (last) step.
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -1- SIMILARITY: Belongs to the trpA family.
DR EMBL; AF173835; AAD51339.1; -.
DR HSSP; P00929; 2WSY.
DR GO; GO:0016829; P:lyase activity; IEA.
DR GO; GO:0004834; P:tryptophan synthase activity; IEA.
DR GO; GO:0000162; P:tryptophan biosynthesis; IEA.
DR InterPro; IPR011060; RibP_bind_barrel.
DR InterPro; IPR002028; Trp_synthaseA.
DR ProDom; PD001535; Trp_synthaseA; 1.
DR TIGRfams; TIGR00262; trpA; 1.
DR PROSITE; PS00167; TRP_SYNTHASE_ALPHA; 1.
KW Amino-acid biosynthesis; Aromatic amino acid biosynthesis; Lyase;
KW Tryptophan biosynthesis.
SQ SEQUENCE 274 AA; 28840 MW; A1971BB93D29A324 CRC64;

Query Match 90.9%; Score 30; DB 2; Length 274;
Best Local Similarity 85.7%; Pred. No. 67;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLRA 7
Db 65 QRANLRA 71

RESULT 7
Q889G6 PRELIMINARY; PRT; 656 AA.
ID Q889G6
AC Q889G6;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE TonB-dependent siderophore receptor, putative.
GN OrderedLocusNames=PSPT00784;
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwinn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
RA Nelson W.C., Daviden T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,
RA Uterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
RA Bender C.L., White O., Fraser C.M., Collier A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186 (2003).
DR EMBL; AE016858; AAO54326.1; -.
DR HSSP; P05825; 1PEP.
DR TIGR; PSPT00784; -.
DR GO; GO:0019857; C:outer membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR010917; TonB_recept_C.
DR PROSITE; PS01156; TONB_DEPENDENT_REC_2; UNKNOWN_1.
KW Complete proteome; Receptor.

```

SQ SEQUENCE 656 AA; 72591 MW; 2238CC4E57A63008 CRC64;
 Query Match 90.9%; Score 30; DB 2; Length 656;
 Best Local Similarity 85.7%; Pred. No. 1.7e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QANLRA 7
 Db 566 QANLRA 572

RESULT 8
 Q6BJ62 PRELIMINARY; PRT; 885 AA.
 AC Q6BJ62
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Similar to CA1375|IPF7158 Candida albicans IPF7158 putative
 DE serine/threonine kinase.
 GN ORFNames=DEHA0G05544g;
 OS Debaryomyces hansenii CBS767.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.
 OX NCBI_TaxID=284592;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=CBS767;
 RG Genolevures;
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
 RA Boisrame A., Boyer J., Catolico L., Confaniolieri F., de Daruvar A.,
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
 RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
 RA Swennene D., Tekai F., Wesolowski-Louvel M., Westhof E., Wirth B.,
 RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
 RA Winkler P., Souciet J.L.;
 RT "Genome evolution in yeasts."
 RN Nature 430:35-44(2004).
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=CBS767;
 RA Genoscope;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; CR382139; CAG90216.1; --
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000209; Pept_S8_S53.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE; PS00138; SUBTILASE SER; UNKNOWN 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 885 AA; 98796 MW; 04A51DF413A54D3E CRC64;

Query Match 90.9%; Score 30; DB 2; Length 885;
 Best Local Similarity 85.7%; Pred. No. 2.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QANLRA 7
 Db 479 QANLRS 485

RESULT 9
 Q7PIB4 PRELIMINARY; PRT; 1018 AA.
 AC Q7PIB4
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE ENSANGP0000024901 (Fragment).
 GN Name=ENSANG00000010544;
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
 OX NCBI_TaxID=180454;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=PEST;
 RA Anopheles Genome Sequencing Consortium;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAB01008960; EAA44209.1; --
 DR GO; GO:0016624; P:oxidoreductase activity, acting on the alde. .; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR001017; Dehydrogenase_E1.
 DR InterPro; IPR005475; Transketolase_CR.
 DR Pfam; PF00676; El_dh; 1.
 DR Pfam; PF02779; Transket_pvr; 1.
 FT NON TER 1018 1018
 SQ SEQUENCE 1018 AA; 114451 MW; F484BB19B8B93298 CRC64;

Query Match 90.9%; Score 30; DB 2; Length 1018;
 Best Local Similarity 85.7%; Pred. No. 2.8e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QANLRA 7
 Db 167 ERANLRA 173

RESULT 10
 Q647G8 PRELIMINARY; PRT; 158 AA.
 AC Q647G8
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE 2S protein 2.
 OS Arachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Arachis.
 OX NCBI_TaxID=3818;
 [1]
 RN SEQUENCE FROM N.A.
 RA Yan Y., Wang L., Huang S.;
 RT "cDNA clone of peanut seed storage protein gene."
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY72691; AAU21496.1; --
 SQ SEQUENCE 158 AA; 18455 MW; 32AD4D4061544948 CRC64;

Query Match 87.9%; Score 29; DB 2; Length 158;
 Best Local Similarity 100.0%; Pred. No. 65;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 6
Db 49 QRANLR 54

RESULT 11

ID Q9SQH1 PRELIMINARY; PRT; 160 AA.
AC Q9SQH1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Allergen.
GN Name=Arachis hypogaea (Peanut).
OS Arachis hypogaea; Streptophyta; Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Virginia; TISSUE=Seed;
RX MEDLINE=99408463; PubMed=10474031; DOI=10.1159/000024203;
RA Kleber-Janke T., Cramer R., Appenzeller U., Schlaak M., Becker W.M.;
RT "Selective cloning of peanut allergens, including profilin and 2S
albumins, by phase display technology";
RL Int. Arch. Allergy Immunol. 119:265-274 (1999).
DR EMBL; AF091737; RAD56719.1; -;
DR InterPro; IPR003612; AAI.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 160 AA; 18417 MW; 9F9E4CEE6808D4C CRC64;

Query Match 87.9%; Score 29; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 6
Db 50 QRANLR 55

RESULT 12

ID Q8PQY7 PRELIMINARY; PRT; 245 AA.
AC Q8PQY7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ABC transporter ATP-binding protein.
GN Name=vehX; OrderedLocusNames=XAC0182;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.B., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,
RA Canarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,

RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463 (2002).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AE011642; AAM35074.1; -;
DR HSP; Q9YGA6; IG29.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC transporter.
DR Pfam; PF00005; ABC tran; 1.
DR ProDom; PD000006; ABC transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 245 AA; 26698 MW; B9BC33B69C48D52D CRC64;

Query Match 87.9%; Score 29; DB 2; Length 245;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 7
Db 66 QRANLR 72

RESULT 13

ID Q63ITO PRELIMINARY; PRT; 251 AA.
AC Q63ITO;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORENAMES=BPSS1986;
OS Burkholderia pseudomallei K96243.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=272560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K96243;
RX PubMed=15377794;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdeno-Tarraga A.M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosset B., Davis P., Deshazer D.,
RA Fellwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitsch E., Rutherford K., Sanders M., Simmonds M.,
RA Songsvilai S., Stevens K., Tumapa S., Vesaratchavest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245 (2004).
DR EMBL; BX571966; CAH39464.1; -;
KW Hypothetical protein.
SQ SEQUENCE 251 AA; 27637 MW; C84954DD87784242 CRC64;

Query Match 87.9%; Score 29; DB 2; Length 251;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 7
Db :|||||

```

Db      136 RANLRA 142
RESULT 14
TRPA_BUCBP
ID      TRPA_BUCBP      STANDARD;      PRT;      269 AA.
AC      P59457;
DT      10-OCT-2003 (Rel. 42, Created)
DT      10-OCT-2003 (Rel. 42, Last sequence update)
DT      25-OCT-2004 (Rel. 45, Last annotation update)
DE      Tryptophan synthase alpha chain (EC 4.2.1.20).
OS      Buchnera aphidicola (subsp. Baizongia pistaciae).
GN      Name=trpa; OrderedLocusNames=bbp257;
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Buchnera.
OX      NCBI_TaxID=135842;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      PubMed=12522265; DOI=10.1073/pnas.0235981100;
RA      van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
RA      Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,
RA      Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;
RT      "Reductive genome evolution in Buchnera aphidicola.";
RL      Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
CC      -!- FUNCTION: The alpha subunit is responsible for the aldol cleavage
CC      of indoleglycerol phosphate to indole and glyceraldehyde 3-
CC      phosphate.
CC      -!- CATALYTIC ACTIVITY: L-serine + 1-(indol-3-yl)glycerol 3-phosphate
CC      = L-tryptophan + glyceraldehyde 3-phosphate.
CC      -!- PATHWAY: Tryptophan biosynthesis; fifth (last) step.
CC      -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC      similarity).
CC      -!- SIMILARITY: Belongs to the trpa family.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; AE014016; AAO26984.1; -
CC      HSSP; P00929; 1K7E.
CC
CC      HAWAP; MF 00131; -.
CC      InterPro; IPR003009; FMN enzyme.
CC      InterPro; IPR01060; RibP_Bind_barrel.
CC      InterPro; IPR002028; Trp_synthaseA.
CC      Pfam; PF00290; Trp_synthA; 1.
CC      ProDom; PD001535; Trp_synthaseA; 1.
CC      TIGRFAMs; TIGR00262; Trpa; 1.
CC      PROSITE; PS00167; TRP_SYNTHASE_ALPHA; 1.
CC      Complete proteome; Lyase; Tryptophan biosynthesis.
KW      ACT SITE 50 Proton acceptor (By similarity).
FT      ACT SITE 50 Proton acceptor (By similarity).
SQ      ACT SITE 61 Proton acceptor (By similarity).
SQ      SEQUENCE 269 AA; 30078 MW; 42E4AA359E8162E8 CRC64;

Query Match      87.9%; Score 29; DB 1; Length 269;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy      1 QANLRA 7
      |:|||||
Db      66 QANLRA 72

RESULT 15
Q8D767
ID      Q8D767      PRELIMINARY;      PRT;      280 AA.
AC      Q8D767;
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE      Outer membrane receptor protein.
DE      OrderedLocusNames=VV20298;
GN      Vibrio vulnificus.
OS      Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC      Vibrionaceae; Vibrio.
OX      NCBI_TaxID=672;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CMCP6;
RA      Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA      Choy H.E.;
RT      "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL      Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AE016809; AAC07260.1; -.
KW      Complete proteome.
SQ      SEQUENCE 280 AA; 31390 MW; 1E2C85898FE1C7A4 CRC64;

Query Match      87.9%; Score 29; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 QANLRA 6
      |:|||||
Db      25 QANLRA 30

Search completed: April 21, 2005, 13:16:43
Job time : 62.8 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2005, 12:51:51 ; Search time 13.1 Seconds
(without alignments)
51.414 Million cell updates/sec

Title: US-10-080-100-46

Perfect score: 33

Sequence: 1 QRANLRA 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: pir1.*
- 2: pir2.*
- 3: pir3.*
- 4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	84.8	268	1	TSEBAT
2	28	84.8	268	2	AB0653
3	28	84.8	399	2	AD3512
4	28	84.8	483	2	G75392
5	28	84.8	502	2	A23547
6	27	81.8	269	2	C84962
7	27	81.8	318	2	T15869
8	27	81.8	390	2	AC3557
9	27	81.8	392	2	S11998
10	27	81.8	613	2	A82834
11	27	81.8	640	2	AG1269
12	27	81.8	640	2	AL1631
13	27	81.8	666	2	B82511
14	27	81.8	1114	2	B86423
15	27	81.8	1358	2	S33653
16	26	78.8	106	2	T34560
17	26	78.8	153	2	B56394
18	26	78.8	230	2	A70459
19	26	78.8	268	2	D82232
20	26	78.8	275	2	H83325
21	26	78.8	282	2	D69379
22	26	78.8	303	2	B95182
23	26	78.8	303	2	E38049
24	26	78.8	334	2	A44894
25	26	78.8	334	2	I40040
26	26	78.8	335	2	I40159
27	26	78.8	336	1	FLYB3
28	26	78.8	336	2	I40088
29	26	78.8	336	2	S37728

30	26	78.8	336	2	I40079	flagellin - Lyme d
31	26	78.8	336	2	A38450	flagellin - Lyme d
32	26	78.8	336	2	S70255	flagellin - Lyme d
33	26	78.8	336	2	S70256	flagellin - Lyme d
34	26	78.8	336	2	I40134	flagellin - Lyme d
35	26	78.8	336	2	I40140	flagellin - Lyme d
36	26	78.8	336	2	I40092	flagellar filament
37	26	78.8	336	2	I40077	flagellin - Lyme d
38	26	78.8	336	2	I40075	flagellin - Lyme d
39	26	78.8	336	2	I40135	flagellin - Lyme d
40	26	78.8	336	2	I40041	flagellin - Borrel
41	26	78.8	336	2	I40204	flagellin - Borrel
42	26	78.8	355	2	T40308	probable sec14 cyt
43	26	78.8	360	2	T47170	hypothetical prote
44	26	78.8	366	2	T42377	hypothetical prote
45	26	78.8	382	2	AG0188	histidinol-phospha

ALIGNMENTS

RESULT 1

TSEBAT

tryptophan synthase (EC 4.2.1.20) alpha chain - Salmonella typhimurium

C:Species: Salmonella typhimurium

C:Date: 28-Feb-1980 #sequence_revision 28-Feb-1980 #text_change 09-Jul-2004

C:Accession: A93837; A92126; A01152

R:Nichols, B.P.; Yanofsky, C.

Proc. Natl. Acad. Sci. U.S.A. 76, 5244-5248, 1979

A:Title: Nucleotide sequences of trpA of Salmonella typhimurium and Escherichia coli: an

A:Reference number: A93837; MUID:80056671; PMID:388433

A:Accession: A93837

A:Molecule type: DNA

A:Residues: 1-268 <NIC>

A:Cross-references: UNIPROT:P00929; GB:V01376; NID:g47938; PIDN:CAA24666.1; PID:g47940

R:Li, S.L.; Yanofsky, C.

J. Biol. Chem. 248, 1830-1836, 1973

A:Title: Amino acid sequence studies with the tryptophan synthetase alpha chain of Salmo

A:Reference number: A92126; MUID:73149276; PMID:4571777

A:Contents: tentative sequence

A:Accession: A92126

A:Molecule type: protein

A:Residues: 1-268 <LIS>

C:Genetics:

A:Gene: trpA

C:Complex: heterotetramer; two alpha and two beta chains

C:Function:

A:Description: catalyzes conversion of indoleglycerol phosphate and serine to tryptophan

A:Pathway: tryptophan biosynthesis

A:Note: cofactor pyridoxal phosphate

A:Note: last step in pathway

C:Superfamily: tryptophan synthase alpha chain; tryptophan synthase alpha chain homology

C:Keywords: carbon-oxygen lyase; heterotetramer; hydro-lyase; tryptophan biosynthesis

F:18-246/Domain: tryptophan synthase alpha chain homology <TRPA>

F:49/Active site: Glu #status predicted

Query Match 84.8%; Score 28; DB 1; Length 268;

Best Local Similarity 85.7%; Pred. No. 45;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QRANLRA 7

Db 65 QRANLRA 71

RESULT 2

AB0653

tryptophan synthase alpha chain [imported] - Salmonella enterica subsp. enterica serovar

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C:Accession: AB0653

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AB0653
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-268 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD08405.1; PID:g16502448; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY1324
 C:Superfamily: tryptophan synthase alpha chain; tryptophan synthase alpha chain homology

Query Match 84.8%; Score 28; DB 2; Length 268;
 Best Local Similarity 85.7%; Pred. No. 45;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QANLRA 7
 |||||
 Db 65 QANLRA 71

RESULT 3
 AD3512
 hypothetical protein BMEI10022 [imported] - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
 C:Accession: AD3512
 R:DelVecchio, V.G.; Kaputrai, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
 ; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
 A:Reference number: AD3512; PMID:1175668
 A:Accession: AD3512
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-399 <KUR>
 A:Cross-references: UNIPROT:Q8YDZ8; GB:AE008918; PIDN:AAL53263.1; PID:g17984144; GSPDB:G
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BMEI10022
 A:Map position: II

Query Match 84.8%; Score 28; DB 2; Length 399;
 Best Local Similarity 85.7%; Pred. No. 67;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QANLRA 7
 |||||
 Db 4 QANLRA 10

RESULT 4
 G75392
 glycosyl hydrolase, family 13 - *Deinococcus radiodurans* (strain R1)
 C:Species: *Deinococcus radiodurans*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: G75392
 R:White, O.; Eissen, J.A.; Heidelberger, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: G75392
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-483 <WHI>
 A:Cross-references: UNIPROT:Q9RUB8; GB:AE001991; GB:AE000513; NID:96459223; PIDN:AAFI103
 A:Experimental source: strain R1
 C:Genetics:

A:Gene: DRI472
 A:Map position: 1
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

Query Match 84.8%; Score 28; DB 2; Length 483;
 Best Local Similarity 100.0%; Pred. No. 81;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RANLRA 7
 |||||
 Db 54 RANLRA 59

RESULT 5
 A23547
 keratin, type II cytoskeletal - African clawed frog
 C:Species: *Xenopus laevis* (African clawed frog)
 C:Date: 23-Aug-1987 #sequence_revision 23-Aug-1987 #text_change 09-Jul-2004
 C:Accession: A23547
 R:Franz, J.K.; Franke, W.W.
 Proc. Natl. Acad. Sci. U.S.A. 83, 6475-6479, 1986
 A:Title: Cloning of cDNA and amino acid sequence of a cytokeratin expressed in oocytes o
 A:Reference number: A23547; MUID:86313601; PMID:2428034
 A:Accession: A23547
 A:Molecule type: mRNA
 A:Residues: 1-502 <PRA>
 A:Cross-references: UNIPROT:P08776; GB:M13811; NID:g214555; PIDN:AAA49891.1; PID:g214555
 C:Superfamily: cytoskeletal keratin
 C:Keywords: coiled coil

Query Match 84.8%; Score 28; DB 2; Length 502;
 Best Local Similarity 85.7%; Pred. No. 84;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QANLRA 7
 |||||
 Db 335 QANLRA 341

RESULT 6
 C84962
 tryptophan synthase (EC 4.2.1.20) alpha chain [imported] - *Buchnera* sp. (strain APS)
 C:Species: *Buchnera* sp.
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: C84962
 R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
 Nature 407, 81-86, 2000
 A:Title: Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. A
 A:Reference number: A84930; MUID:20445173; PMID:10993077
 A:Accession: C84962
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-269 <STO>
 A:Cross-references: GB:AP000398; GSPDB:GN00144
 A:Experimental source: strain APS
 C:Genetics:
 A:Gene: trpA; BU277
 C:Superfamily: tryptophan synthase alpha chain; tryptophan synthase alpha chain homology
 C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 81.8%; Score 27; DB 2; Length 269;
 Best Local Similarity 71.4%; Pred. No. 78;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QANLRA 7
 |||||
 Db 65 QANLRA 71

RESULT 7
 T15869
 hypothetical protein C56G2.3 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*


```
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T15869
R;Connell, M.
A;Description: The sequence of C. elegans cosmid C56G2.
A;Reference number: Z18420
A;Accession: T15869
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-318 <CON>
A;Cross-references: UNIPROT:Q09287; EMBL:U23177; NID:g726411; PID:g726415; PIDN:AAA64330
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:C56G2.3
A;introns: 178/1

Query Match      81.8%; Score 27; DB 2; Length 318;
Best Local Similarity 71.4%; Pred. No. 92;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLRA 7
   :|||:
Db 226 RRANIRA 232

RESULT 8
AC3557
acriflavin resistance protein a precursor [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AC3557
R;DeIvecchio, V.G.; Kaputral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A;Reference number: AD3252; PMID:11756688
A;Accession: AC3557
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-390 <KUR>
A;Cross-references: UNIPROT:Q8YCZ7; GB:AE008918; PIDN:AAL53622.1; PID:g17984537; GSPDB:G
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEII0380
A;Map position: II

Query Match      81.8%; Score 27; DB 2; Length 390;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 6
   :|||:
Db 121 QRANWR 126

RESULT 9
S11998
finger protein odd-skipped - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
C;Accession: S11998
R;Coulter, D.E.; Swaykus, E.A.; Beran-Kohn, M.A.; Goldberg, D.; Wieschaus, E.; Schedl,
EMBO J. 8, 3795-3804, 1990
A;Title: Molecular analysis of odd-skipped, a zinc finger encoding segmentation gene with
A;Reference number: S11998
A;Accession: S11998
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-392 <COU>

Query Match      81.8%; Score 27; DB 2; Length 392;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 QRANLRA 7
   :|||:
Db 315 QRANLKS 321

RESULT 10
A82834
hypothetical protein XF0221 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: A82834
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: A82834
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-613 <SIM>
A;Cross-references: UNIPROT:Q9FGS7; GB:AE003875; GB:AE003849; NID:g9105019; PIDN:AAF8303
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.G. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0221

Query Match      81.8%; Score 27; DB 2; Length 613;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QRANLRA 7
   :|||:
Db 240 QRAQLRA 246

RESULT 11
AG1269
threonyl-tRNA synthetase [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AG1269
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kraft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueder, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AG1269
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-640 <GLA>
A;Cross-references: UNIPROT:Q8Y6X2; GB:NC_003210; PIDN:CAC99637.1; PID:gl6410988; GSPDB:
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: thrS
C;Superfamily: threonine-tRNA ligase
```

```

Query Match      81.8%; Score 27; DB 2; Length 640;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QRANLRA 7
    ||| |||
Db 568 QRAGLRA 574

RESULT 12
A11631
threonyl-tRNA synthetase [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: A11631
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: A11631
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-640 <GLA>
A:Cross-references: UNIPROT:Q92BF5; GB:AL592022; PIDN:CAC96825.1; PID:gl6414081; GSPDB:C
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: thrS
C:Superfamily: threonine-tRNA ligase

Query Match      81.8%; Score 27; DB 2; Length 640;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QRANLRA 7
    ||| |||
Db 568 QRAGLRA 574

RESULT 13
D82511
1,4-alpha-glucan branching enzyme VCA0016 [imported] - Vibrio cholerae (strain N16961 se
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: D82511
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathavan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20408833; PMID:10952301
A:Accession: D82511
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-666 <HEI>
A:Cross-references: UNIPROT:Q9KNE8; GB:AB004345; GB:AE003853; NID:g9657390; PIDN:AAF9593
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0016
A:Map position: 2
C:Superfamily: 1,4-alpha-glucan branching enzyme

Query Match      81.8%; Score 27; DB 2; Length 666;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QRANLRA 7
    ||| |||
Db 486 QTANLRA 492

```

```

RESULT 14
B86423
hypothetical protein T1P2.15 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: B86423
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Chung, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Roney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: B86423
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1114 <STO>
A:Cross-references: UNIPROT:Q9C8S4; GB:AE005172; NID:gl0092454; PIDN:AAG12856.1; GSPDB:G
C:Genetics:
A:Map position: 1

Query Match      81.8%; Score 27; DB 2; Length 1114;
Best Local Similarity 83.3%; Pred. No. 3.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 6
    |||||
Db 817 QRANIR 822

RESULT 15
S33653
probable serine/threonine protein kinase (EC 2.7.1.-) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YAL002; protein YAL017w; secretory protein SSP138
C:Species: Saccharomyces cerevisiae
C:Date: 30-Sep-1993 #sequence_revision 02-Aug-1994 #text_change 16-Aug-2004
C:Accession: S33653; S36717; S36732; JH0486
R:Clark, M.W.; Zhong, W.W.; Keng, T.; Storms, R.K.; Ouellette, B.F.F.; Barton, A.; Kabac
Yeast 9, 543-549, 1993
A:Title: The YAL017 gene on the left arm of chromosome I of Saccharomyces cerevisiae enc
A:Reference number: S33653; MUID:93311122; PMID:8322517
A:Accession: S33653
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1358 <CLA>
A:Cross-references: UNIPROT:P31374; EMBL:L05146
R:Ouellette, B.; Clark, M.W.; Keng, T.; Storms, R.K.; Zhong, W.; Zeng, B.; Fortin, N.; D
submitted to the EMBL Data Library, January 1993
A:Description: Sequencing of Chromosome I from Saccharomyces cerevisiae: analysis of a 3
A:Reference number: S36711
A:Accession: S36717
A:Molecule type: DNA
A:Residues: 1-864, 867-1358 <OUB>
A:Cross-references: EMBL:L05146; NID:gl71851; PIDN:AAC04940.1; PID:gl71858; MIPS:YAL017w
R:Clark, M.W.; Zhong, W.W.; Keng, T.; Storms, R.K.; Barton, A.; Kaback, D.B.; Bussey, H.
Yeast 8, 133-145, 1992
A:Title: Identification of a Saccharomyces cerevisiae homolog of the SNF2 transcriptiona
A:Reference number: S22266; MUID:92221690; PMID:1561836
A:Accession: S36732
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-862 <CL2>
A:Cross-references: EMBL:S93805
R:Sidhu, R.S.; Mathewes, S.; Bollon, A.P.
Gene 107, 111-118, 1991
A:Title: Selection of secretory protein-encoding genes by fusion with PHO5 in Saccharomy
A:Reference number: JH0483; MUID:92077420; PMID:1743509
A:Accession: JH0486

```

A:Molecule type: DNA
 A:Residues: 1-72,'E',74-154 <SID>
 C:Genetics:
 A:Gene: SGD:FUN31; SSP138
 A:Cross-references: SGD:S0000015; MIPS:YAL017w
 A:Map position: 1L
 C:Superfamily: protein kinase homology
 E:1096-1356/Domain: protein kinase homology <KINS>
 F:1104-1112/Region: protein kinase ATP-binding motif
 F:8,128/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:1232/Active site: Asp #status predicted

Query Match 81.8%; Score 27; DB 2; Length 1358;
 Best Local Similarity 71.4%; Pred. No. 4e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 QRANLRA 7
 |||||:
 Db 334 QRANLKS 340

Search completed: April 21, 2005, 13:06:27
 Job time : 13.2 secs

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2005, 13:06:46 ; Search time 52.7 Seconds
(without alignments)
44.204 Million cell updates/sec

Title: US-10-080-100-46

Perfect score: 33

Sequence: 1 ORANLRA 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1424015 seqs, 332791073 residues

Total number of hits satisfying chosen parameters: 1424015

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	7	13	US-10-080-100-46
2	33	100.0	8	13	Sequence 46, Appl
3	29	87.9	392	16	US-10-080-100-10
4	29	87.9	510	9	US-09-874-923-50
5	29	87.9	510	9	US-09-991-496-50
6	29	87.9	538	9	US-09-874-923-118
7	29	87.9	538	9	US-09-991-496-118
8	29	87.9	735	15	US-10-369-493-19307
9	29	87.9	768	16	US-10-437-963-188571
10	29	87.9	868	16	US-10-437-963-188566
11	29	87.9	869	16	US-10-437-963-188578
12	29	87.9	877	16	US-10-437-963-188584
13	29	87.9	886	16	US-10-437-963-188587

14	29	87.9	916	16	US-10-437-963-188581
15	29	87.9	1822	15	US-10-432-443-39
16	28	84.8	1822	15	US-10-424-599-206747
17	28	84.8	120	15	US-10-424-599-224070
18	28	84.8	166	9	US-09-731-221-77
19	28	84.8	170	15	US-10-425-114-66342
20	28	84.8	253	16	US-10-767-701-41634
21	28	84.8	268	16	US-10-872-198-51
22	28	84.8	268	17	US-10-872-197A-51
23	28	84.8	391	16	US-10-437-963-202800
24	28	84.8	439	16	US-10-788-792-239
25	28	84.8	440	14	US-10-221-873-4
26	28	84.8	441	16	US-10-408-765A-1563
27	28	84.8	442	10	US-09-815-379-16
28	28	84.8	447	15	US-10-363-616-392
29	28	84.8	453	14	US-10-106-698-5604
30	28	84.8	483	15	US-10-369-493-597
31	28	84.8	498	16	US-10-605-708A-2
32	28	84.8	873	16	US-10-437-963-193309
33	28	84.8	1572	17	US-10-487-132-5
34	27	81.8	76	9	US-09-843-676-205
35	27	81.8	76	10	US-09-438-486-205
36	27	81.8	76	14	US-10-053-758-205
37	27	81.8	76	14	US-10-054-295-205
38	27	81.8	76	14	US-10-054-611-205
39	27	81.8	76	15	US-10-325-810-324
40	27	81.8	76	17	US-10-877-146-324
41	27	81.8	120	15	US-10-424-599-184173
42	27	81.8	126	16	US-10-437-963-170215
43	27	81.8	175	15	US-10-424-599-159578
44	27	81.8	176	15	US-10-424-599-263552
45	27	81.8	178	15	US-10-424-599-168045

ALIGNMENTS

RESULT 1

```
US-10-080-100-46
; Sequence 46, Application US/10080100
; Publication No. US20020165356A1
; GENERAL INFORMATION:
; APPLICANT: Barbas, Carlos
; APPLICANT: Dreier, Birgit
; TITLE OF INVENTION: Zinc Finger Binding Domains for
; FILE REFERENCE: TSRI 760.0
; CURRENT APPLICATION NUMBER: US/10/080.100
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US/09/791.106
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthesized
US-10-080-100-46
```

Query Match 100.0%; Score 33; DB 13; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ORANLRA 7

Db 1 ORANLRA 7

RESULT 2

```
US-10-080-100-10
; Sequence 10, Application US/10080100
```

```
; Publication No. US20020165356A1
; GENERAL INFORMATION:
; APPLICANT: Barbas, Carlos
; APPLICANT: Dreier, Birgit
; TITLE OF INVENTION: Zinc Finger Binding Domains for
; FILE REFERENCE: TSRI 760.0
; CURRENT APPLICATION NUMBER: US/10/080,100
; PRIOR FILING DATE: 2002-02-21
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthesized
US-10-080-100-10

Query Match      100.0%; Score 33; DB 13; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QRANLRA 7
      |||||
Db      2 QRANLRA 8

RESULT 3
US-10-437-963-188574
; Sequence 188574, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188574
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85165C.1.pep
US-10-437-963-188574

Query Match      87.9%; Score 29; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QRANLRA 6
      |||||
Db      153 QRANLRA 158

RESULT 4
US-09-874-923-50
; Sequence 50, Application US/09874923
; Patent No. US20020081320A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
```

```
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C8
; CURRENT APPLICATION NUMBER: US/09/874,923
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Leishmania chagasi
US-09-874-923-50

Query Match      87.9%; Score 29; DB 9; Length 510;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QRANLR 6
      |||||
Db      114 QRANLR 119

RESULT 5
US-09-991-496-50
; Sequence 50, Application US/09991496
; Patent No. US20020169285A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C9
; CURRENT APPLICATION NUMBER: US/09/991,496
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Leishmania chagasi
US-09-991-496-50

Query Match      87.9%; Score 29; DB 9; Length 510;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QRANLR 6
      |||||
Db      114 QRANLR 119

RESULT 6
US-09-874-923-118
; Sequence 118, Application US/09874923
; Patent No. US20020081320A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
```

; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C8
; CURRENT APPLICATION NUMBER: US/09/874,923
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 118
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Leishmania major and chagasi
US-09-874-923-118

Query Match 87.9%; Score 29; DB 9; Length 538;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 6
Db 117 QRANLR 122

RESULT 7
US-09-991-496-118
; Sequence 118, Application US/09991496
; Patent No. US20020169285A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C9
; CURRENT APPLICATION NUMBER: US/09/991,496
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 118
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Leishmania major and chagasi
US-09-991-496-118

Query Match 87.9%; Score 29; DB 9; Length 538;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 6
Db 117 QRANLR 122

RESULT 8
US-10-369-493-19307
; Sequence 19307, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19307
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-10-369-493-19307

Query Match 87.9%; Score 29; DB 15; Length 735;
Best Local Similarity 85.7%; Pred. No. 5.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 7
Db 557 KRANLRA 563

RESULT 9
US-10-437-963-188571
; Sequence 188571, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188571
; LENGTH: 768
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85162C.1.pep
US-10-437-963-188571

Query Match 87.9%; Score 29; DB 16; Length 768;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 6
Db 228 QRANLR 233

RESULT 10
US-10-437-963-188566
; Sequence 188566, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188566
; LENGTH: 868
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85158C.1.pep
US-10-437-963-188566

Query Match 87.9%; Score 29; DB 16; Length 868;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 Qranlr 6
Db 201 Qranlr 206

RESULT 11
US-10-437-963-188578
; Sequence 188578, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188578
; LENGTH: 869
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85169C.1.pep
US-10-437-963-188578

Query Match 87.9%; Score 29; DB 16; Length 869;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 Qranlr 6
Db 249 Qranlr 254

RESULT 12
US-10-437-963-188584
; Sequence 188584, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188584
; LENGTH: 877
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85174C.1.pep
US-10-437-963-188584

Query Match 87.9%; Score 29; DB 16; Length 877;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 Qranlr 6
Db 240 Qranlr 245

RESULT 13
US-10-437-963-188587
; Sequence 188587, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188587
; LENGTH: 886
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85177C.1.pep
US-10-437-963-188587

Query Match 87.9%; Score 29; DB 16; Length 886;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 Qranlr 6
Db 244 Qranlr 249

RESULT 14
US-10-437-963-188581
; Sequence 188581, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188581
; LENGTH: 916
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85171C.1.pap
US-10-437-963-188581

Query Match 87.9%; Score 29; DB 16; Length 916;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 6
Db 249 QRANLR 254

RESULT 15
US-10-432-443-39
; Sequence 39, Application US/10432443
; Publication No. US20040076982A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn et al.
; TITLE OF INVENTION: 3-HYDROXYPROPIONIC ACID AND OTHER ORGANIC COMPOUNDS
; FILE REFERENCE: 6682-65884
; CURRENT APPLICATION NUMBER: US/10/432,443
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/43607
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: US 60/317,845
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/306,727
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/285,478
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/252,123
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 188
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 1822
; TYPE: PRT
; ORGANISM: Chloroflexus aurantiacus
US-10-432-443-39

Query Match 87.9%; Score 29; DB 15; Length 1822;
Best Local Similarity 85.7%; Pred. No. 1.4e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLRA 7
Db 1554 QRANLRA 1560

Search completed: April 21, 2005, 14:09:43
Job time : 54.7 secs

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2005, 12:51:52 ; Search time 18.7 Seconds
(without alignments)
27.943 Million cell updates/sec

Title: US-10-080-100-46

Perfect score: 33

Sequence: 1 QRANLRA 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	87.9	177	4	US-09-252-991A-31771
2	29	87.9	510	3	US-09-183-861-50
3	29	87.9	510	3	US-09-022-785-50
4	29	87.9	510	4	US-09-551-974A-50
5	29	87.9	510	4	US-09-565-501A-50
6	29	87.9	510	4	US-09-639-206A-50
7	29	87.9	510	4	US-09-874-923-50
8	29	87.9	538	4	US-09-874-923-118
9	29	87.9	625	4	US-09-902-540-12183
10	28	84.8	104	4	US-09-621-976-4383
11	28	84.8	124	4	US-09-513-999C-5279
12	28	84.8	307	4	US-09-902-540-10505
13	28	84.8	430	4	US-09-902-540-13985
14	28	84.8	457	4	US-09-252-991A-29671
15	28	84.8	673	3	US-09-075-272-5
16	28	84.8	3038	1	US-08-450-332-2
17	28	84.8	3038	1	US-08-637-640-2
18	28	84.8	3038	3	US-09-004-406C-2
19	27	81.8	76	3	US-08-851-843A-205
20	27	81.8	76	3	US-08-974-549A-324
21	27	81.8	76	3	US-08-854-050-205
22	27	81.8	76	3	US-09-430-323-205
23	27	81.8	76	4	US-09-402-181B-324
24	27	81.8	76	4	US-09-721-456-324
25	27	81.8	92	4	US-09-248-796A-16168
26	27	81.8	240	4	US-09-902-540-15786
27	27	81.8	301	4	US-09-252-991A-18428

28	27	81.8	327	4	US-09-270-767-42278	Sequence 42278, A
29	27	81.8	332	4	US-09-252-991A-20871	Sequence 20871, A
30	27	81.8	472	4	US-09-166-350-17	Sequence 17, Appl
31	27	81.8	654	4	US-09-252-991A-19805	Sequence 19805, A
32	27	81.8	1356	3	US-09-770-170-6	Sequence 6, Appl
33	26	78.8	28	4	US-09-191-593-53	Sequence 53, Appl
34	26	78.8	28	4	US-09-191-593-58	Sequence 58, Appl
35	26	78.8	82	4	US-09-270-767-38253	Sequence 38253, A
36	26	78.8	82	4	US-09-270-767-53470	Sequence 53470, A
37	26	78.8	108	1	US-08-166-160-19	Sequence 19, Appl
38	26	78.8	127	4	US-09-270-767-38390	Sequence 38390, A
39	26	78.8	127	4	US-09-270-767-53607	Sequence 53607, A
40	26	78.8	157	4	US-09-106-872A-22	Sequence 22, Appl
41	26	78.8	157	4	US-09-191-593-10	Sequence 10, Appl
42	26	78.8	157	4	US-09-191-593-21	Sequence 21, Appl
43	26	78.8	159	4	US-09-106-872A-2	Sequence 2, Appl
44	26	78.8	213	2	US-08-437-013-2	Sequence 2, Appl
45	26	78.8	213	4	US-09-275-506A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-09-252-991A-31771
; Sequence 31771, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 31142
; SEQ ID NO 31771
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31771

Query Match 87.9%; Score 29; DB 4; Length 177;
Best Local Similarity 100.0%; Pred.No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 6
Db 112 QRANLR 117

RESULT 2

US-09-183-861-50
; Sequence 50, Application US/09183861
; Patent No. 6365165
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USEIN THE THERAPY AND
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/183.861
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/022,765
FILING DATE: 12-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.420C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 510 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Leishmania chagasi
US-09-183-861-50

Query Match 87.9%; Score 29; DB 3; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 6
Db 114 QRANLR 119

RESULT 3
US-09-022-765-50
; Sequence 50, Application US/09022765
; Patent No. 6375955
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USEIN THE THERAPY AND
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,765
FILING DATE: 12-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.420C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 510 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: linear
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: Leishmania chagasi
US-09-022-765-50

Query Match 87.9%; Score 29; DB 3; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 6
Db 114 QRANLR 119

RESULT 4
US-09-551-974A-50
; Sequence 50, Application US/09551974A
; Patent No. 6500437
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C5
; CURRENT APPLICATION NUMBER: US/09/551,974A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Leishmania chagasi
US-09-551-974A-50

Query Match 87.9%; Score 29; DB 4; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 6
Db 114 QRANLR 119

RESULT 5
US-09-565-501A-50
; Sequence 50, Application US/09565501A
; Patent No. 6607731
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Peter Probst
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C6
; CURRENT APPLICATION NUMBER: US/09/565,501A
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSEQ for Windows Version 4.0

```
; SEQ ID NO 50
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Leishmania chagasi
US-09-565-501A-50

Query Match      87.9%; Score 29; DB 4; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QANLR 6
Db      114 QANLR 119

RESULT 6
US-09-639-206A-50
; Sequence 50, Application US/09639206A
; Patent No. 6613337
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C7
; CURRENT APPLICATION NUMBER: US/09/639,206A
; CURRENT FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Leishmania chagasi
US-09-639-206A-50

Query Match      87.9%; Score 29; DB 4; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QANLR 6
Db      114 QANLR 119

RESULT 7
US-09-874-923-50
; Sequence 50, Application US/09874923
; Patent No. 6638517
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C8
; CURRENT APPLICATION NUMBER: US/09/874,923
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Leishmania chagasi
US-09-874-923-50

Query Match      87.9%; Score 29; DB 4; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QANLR 6
Db      114 QANLR 119

RESULT 8
US-09-874-923-118
; Sequence 118, Application US/09874923
; Patent No. 6638517
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C8
; CURRENT APPLICATION NUMBER: US/09/874,923
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 118
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Leishmania major and chagasi
US-09-874-923-118

Query Match      87.9%; Score 29; DB 4; Length 538;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QANLR 6
Db      117 QANLR 122

RESULT 9
US-09-902-540-12183
; Sequence 12183, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 39-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12183
; LENGTH: 625
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12183

Query Match      87.9%; Score 29; DB 4; Length 625;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
```

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLRA 7

Db 446 KRALRA 452

RESULT 10

US-09-621-976-4383
; Sequence 4383, Application US/09621976
; Patent No. 6839063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Giordano, J.Y.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4383
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-4383

Query Match 84.8%; Score 28; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RANLRA 7

Db 27 RANLRA 32

RESULT 11

US-09-513-999C-5279
; Sequence 5279, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513.999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5279
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 109
; OTHER INFORMATION: Xaa=Arg or Ser
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 113
; OTHER INFORMATION: Xaa= * or Cys or Phe or Leu or Ser or Trp or Tyr
US-09-513-999C-5279

Query Match 84.8%; Score 28; DB 4; Length 124;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RANLRA 7

|||||

Db 27 RANLRA 32

RESULT 12

US-09-902-540-10505
; Sequence 10505, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902.540
; CURRENT FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10505
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10505

Query Match 84.8%; Score 28; DB 4; Length 307;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLRA 7

Db 263 QRANIRS 269

RESULT 13

US-09-902-540-13985
; Sequence 13985, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902.540
; CURRENT FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13985
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13985

Query Match 84.8%; Score 28; DB 4; Length 430;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RANLRA 7

Db 339 RANLRA 344

RESULT 14

US-09-252-991A-29671
; Sequence 29671, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 29671
LENGTH: 457
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29671

Query Match 84.8%; Score 28; DB 4; Length 457;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 QANLRA 7
|||
Db 201 QANLRA 207
|||

RESULT 15

US-09-075-272-5
Sequence 5, Application US/09075272
Patent No. 6136598
GENERAL INFORMATION:
APPLICANT: MILLER, A. DUSTY
APPLICANT: WOLGAMOT, GREG
APPLICANT: BONHAM, LYNN
TITLE OF INVENTION: MUS DUNNI ENDOGENOUS RETROVIRAL
TITLE OF INVENTION: PACKAGING CELL LINES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: U.S.A.
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/075,272
FILING DATE: 08-MAY-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/046,140
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: POOR, Brian W.
REGISTRATION NUMBER: 32,928
REFERENCE/DOCKET NUMBER: 14538A-003710
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 467-9600
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 673 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-075-272-5

Query Match 84.8%; Score 28; DB 3; Length 673;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 RANLRA 7

Db 120 RANLRA 125
|||||

Search completed: April 21, 2005, 13:19:54
Job time : 19.7 secs

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2005, 12:51:51 ; Search time 73.1 Seconds
(without alignments)
37.036 Million cell updates/sec

Title: US-10-080-100-46

Perfect score: 33

Sequence: 1 QRANLRA 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	7	5 ABP53146	Abp53146 Zinc fing
2	33	100.0	7	5 ABP53144	Abp53144 Zinc fing
3	33	100.0	7	6 ABU60749	Abu60749 Phage dis
4	33	100.0	7	6 ABU60796	Abu60796 Phage dis
5	33	100.0	8	5 ABP53167	Abp53167 Zinc fing
6	33	100.0	8	6 ABU60713	Abu60713 Phage dis
7	33	100.0	8	7 ADJ98368	Adj98368 Zinc fing
8	33	100.0	600	7 ADK14951	Adk14951 Urinary s
9	29	87.9	177	7 ABO83025	Abc83025 Pseudomon
10	29	87.9	510	2 AAW70232	Aaw70232 Leishmani
11	29	87.9	510	5 AAE24946	Aae24946 Leishmani
12	29	87.9	510	5 AAU71828	Aau71828 Leishmani
13	29	87.9	510	5 ABG60884	Abg60884 Leishmani
14	29	87.9	510	5 AAB71282	Aab71282 L. chagas
15	29	87.9	510	7 ADB78813	Adb78813 Leishmani
16	29	87.9	538	5 AAB71317	Aab71317 L. major
17	29	87.9	538	7 ADB78881	Adb78881 Leishmani
18	29	87.9	735	8 ADS30274	Ads30274 Bacterial
19	29	87.9	1822	5 ABG80407	Abg80407 3-hydroxy
20	28	84.8	15	3 AAB33582	Aab33582 5/15 nati
21	28	84.8	15	3 AAB33581	Aab33581 5/15 nati
22	28	84.8	15	3 AAB23132	Aab23132 Modified
23	28	84.8	15	3 AAB23133	Aab23133 Modified
24	28	84.8	15	4 AAU04795	Aau04795 Modified
25	28	84.8	15	4 AAU04796	Aau04796 Modified

26	28	84.8	124	3	AAG01198	Aag01198 Human sec
27	28	84.8	157	6	ABU52577	Abu52577 Peanut Ar
28	28	84.8	157	6	ABU52575	Abu52575 Peanut Ar
29	28	84.8	166	3	AAB33600	Aab33600 Modified
30	28	84.8	166	3	AAU05035	Aau05035 Modified
31	28	84.8	167	4	AAU04710	Aau04710 Modified
32	28	84.8	175	8	ADQ20104	Adq20104 Human sof
33	28	84.8	269	6	ABM67453	Abm67453 Photorhab
34	28	84.8	439	8	ADR99233	Adr99233 Nucleolar
35	28	84.8	440	4	AAG79110	Aag79110 Amino aci
36	28	84.8	440	4	AAB94643	Aab94643 Human pro
37	28	84.8	440	8	ADN04942	Adn04942 Antipsori
38	28	84.8	440	8	ABM81606	Abm81606 Tumour-as
39	28	84.8	441	5	AAM52655	Aam52655 Human DNA
40	28	84.8	441	7	ADJ69757	Adj69757 Human hea
41	28	84.8	442	4	AAE11893	Aae11893 Angiogene
42	28	84.8	447	5	ABP62955	Abp62955 Human pol
43	28	84.8	453	4	AAG74830	Aag74830 Human COL
44	28	84.8	457	7	ABO80925	Abc80925 Pseudomon
45	28	84.8	483	8	ADN17944	Adn17944 Bacterial

ALIGNMENTS

RESULT 1

ABP53146

ID ABP53146 standard; peptide; 7 AA.

XX AC ABP53146;

XX AC ABP53146;

DT 12-NOV-2002 (first entry)

XX DE

DE Zinc finger nucleotide binding peptide SEQ ID NO:98.

XX KW

KW Zinc finger binding domain; zinc finger nucleotide binding peptide;

KW zinc finger; gene expression; modulation; promoter; viral; lentivirus;

KW human T-cell lymphocytic virus; HTLV; human immunodeficiency virus; HIV.

XX OS

OS Synthetic.

XX PN

PN WO200266640-A2.

XX PD

PD 29-AUG-2002.

XX PF

PF 21-FEB-2002; 2002WO-EP001862.

XX PR

PR 21-FEB-2001; 2001US-00791106.

XX PA

PA (NOVS) NOVARTIS AG.

PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.

XX PA

PA (SCRI) SCRIPPS RES INST.

XX PI

PI Barbas CF, Dreier B;

XX DR

DR WPI; 2002-674941/72.

XX PT

PT New polypeptides comprising 2 to 10 zinc finger-nucleotide binding

PT peptides, useful for modulating gene expression in promoters from viral

PT groups, e.g. Lentivirus group, such HTLV-1 and 2, or HIV-1 and 2.

XX PS

PS Disclosure; Page 19; 48pp; English.

XX CC

CC The present invention describes a polypeptide comprising 2 to 12 zinc

CC finger-nucleotide binding peptides in which at least one contains a

CC nucleotide binding region. Also described: (1) an isolated and purified

CC polynucleotide encoding the polypeptide cited above; (2) an expression

CC vector containing the polynucleotide in (1); and (3) a process of

CC regulating expression of a nucleotide sequence that contains the sequence

CC S1 comprising exposing the nucleotide sequence to the polypeptide cited

CC above. (S1) is (5'-ANN)n-3', where n = any integer from 2-12; and N = A,

CC C, G, or T. The polypeptide is useful for modulating gene expression in

CC promoters from viral groups, e.g. Lentivirus group, such as human T-cell

CC lymphocytic virus (HTLV) 1 and 2, or human immunodeficiency virus (HIV) 1
 CC or 2. ABQ75687 to ABQ75697 and ABP53112 to ABP53221 represent sequences
 CC given in the exemplification of the present invention
 XX
 SQ Sequence 7 AA;

Query Match 100.0%; Score 33; DB 5; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLRA 7
 |||||
 Db 1 QRANLRA 7

RESULT 2
 ABP53144
 ID ABP53144 standard; peptide; 7 AA.
 AC
 XX ABP53144;
 XX

DT 12-NOV-2002 (first entry)

XX Zinc finger nucleotide binding peptide SEQ ID NO:46.

XX Zinc finger binding domain; zinc finger nucleotide binding peptide;
 KW zinc finger; gene expression; modulation; promoter; viral; Lentivirus;
 KW human T-cell lymphocytic virus; HTLV; human immunodeficiency virus; HIV.
 XX

OS Synthetic.

XX WO200266640-A2.

XX 29-AUG-2002.

XX 21-FEB-2002; 2002WO-EP001862.

XX 21-FEB-2001; 2001US-00791106.

XX (NOVS) NOVARTIS AG

PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.

PA (SCRI) SCRIPPS RES INST.

XX Barbas CF, Dreier B;

XX WPI; 2002-674941/72.

XX New polypeptides comprising 2 to 10 zinc finger-nucleotide binding
 PT peptides, useful for modulating gene expression in promoters from viral
 PT groups, e.g. Lentivirus group, such as HTLV-1 and 2, or HIV-1 and 2.

XX Claim 1; Page 19; 48pp; English.

XX The present invention describes a polypeptide comprising 2 to 12 zinc
 CC finger-nucleotide binding peptides in which at least one contains a
 CC nucleotide binding region. Also described: (1) an isolated and purified
 CC polynucleotide encoding the polypeptide cited above; (2) an expression
 CC vector containing the polynucleotide in (1); and (3) a process of
 CC regulating expression of a nucleotide sequence that contains the sequence
 CC S1 comprising exposing the nucleotide sequence to the polypeptide cited
 CC above. (S1) is (5'-ANN)n-3', where n = any integer from 2-12; and N = A,
 CC C, G, or T. The polypeptide is useful for modulating gene expression in
 CC promoters from viral groups, e.g. Lentivirus group, such as human T-cell
 CC lymphocytic virus (HTLV) 1 and 2, or human immunodeficiency virus (HIV) 1
 CC or 2. ABQ75687 to ABQ75697 and ABP53112 to ABP53221 represent sequences
 CC given in the exemplification of the present invention
 XX

XX Sequence 7 AA;

Query Match 100.0%; Score 33; DB 5; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLRA 7
 |||||
 Db 1 QRANLRA 7

RESULT 3
 ABU60749

ID ABU60749 standard; peptide; 7 AA.

XX AC ABU60749;

XX 06-MAY-2003 (first entry)

XX Phage displayed zinc finger recognising AAA #5.

XX Zinc finger; phage display; ant-HIV; virucide; HIV infection;
 KW human T cell lymphotropic virus infection; HTLV.

XX Synthetic.

XX US2002165356-A1.

XX 07-NOV-2002.

XX 21-FEB-2002; 2002US-00080100.

XX 21-FEB-2001; 2001US-00367356.

XX (SCRI) SCRIPPS RES INST.

XX Barbas CF, Dreier B;

XX WPI; 2003-255225/25.

XX New zinc finger-nucleotide binding polypeptides, useful in the treatment
 PT of disorders associated with the aberrant expression or activity of the
 PT polypeptide, such as HIV infection or human T cell lymphotropic virus
 PT infection.

XX Claim 1; Fig 3; 22pp; English.

XX The invention relates a new polypeptide comprising from 2-12 zinc finger-
 CC nucleotide binding peptides at least one of which contains a nucleotide
 CC binding region having any of 71 nucleotide binding sequences, fully
 CC defined in the specification. Also included are an isolated and purified
 CC polynucleotide that encodes the polypeptide, an expression vector
 CC containing the polynucleotide and a process of regulating expression of a
 CC nucleotide sequence that contains the sequence (5'-ANN)n-3', where n is
 CC an integer from 2-12, the process comprising exposing the nucleotide
 CC sequence to the zinc finger protein of the invention. The methods and
 CC compositions of the present invention, are useful for modulating zinc-
 CC finger protein binding and hence regulating aspects of gene expression.
 CC The zinc finger polypeptides are useful in the treatment of treatment of
 CC disorders associated with the aberrant expression or activity of the
 CC polypeptide, such as viral infection, in particular HIV or human T cell
 CC lymphotropic virus (HTLV) infection. The present sequence is a synthetic
 CC nucleotide binding sequence (zinc finger) expressed by a phage display
 CC library, which may be assembled into the zinc finger protein of the
 CC invention

XX Sequence 7 AA;

Query Match 100.0%; Score 33; DB 6; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLRA 7
 |||||
 Db 1 QRANLRA 7

RESULT 4
 ABU60796

ID ABU60796 standard; peptide; 7 AA.
 AC ABU60796;
 XX
 DT 06-MAY-2003 (first entry)
 XX
 DE Phage displayed zinc finger recognising AAA #6.
 XX
 KW Zinc finger; phage display; ant-HIV; virucide; HIV infection;
 KW human T cell lymphotropic virus infection; HTLV.
 XX
 OS Synthetic.
 XX
 XX US2002165356-A1.
 PN
 PD 07-NOV-2002.
 XX
 XX 21-FEB-2002; 2002US-00080100.
 XX
 PF 21-FEB-2001; 2001US-00367356.
 PR
 XX (SCRI) SCRIPPS RES INST.
 PA
 XX Barbas CF, Dreier B;
 XX WPI; 2003-255225/25.
 XX
 PT New zinc finger-nucleotide binding polypeptides, useful in the treatment
 of disorders associated with the aberrant expression or activity of the
 PT polypeptide, such as HIV infection or human T cell lymphotropic virus
 PT infection.
 XX
 PS Disclosure; Page 7; 22pp; English.
 XX
 CC The invention relates a new polypeptide comprising from 2-12 zinc finger-
 CC nucleotide binding peptides at least one of which contains a nucleotide
 CC binding region having any of 71 nucleotide binding sequences, fully
 CC defined in the specification. Also included are an isolated and purified
 CC polynucleotide that encodes the polypeptide, an expression vector
 CC containing the polynucleotide and a process of regulating expression of a
 CC nucleotide sequence that contains the sequence (5'-ANN) n-3', where n is
 CC an integer from 2-12, the process comprising exposing the nucleotide
 CC sequence from 2-12, the process comprising exposing the nucleotide
 CC sequence to the zinc finger protein of the invention. The methods and
 CC compositions of the present invention, are useful for modulating zinc-
 CC finger protein binding and hence regulating aspects of gene expression.
 CC The zinc finger polypeptides are useful in the treatment of treatment of
 CC disorders associated with the aberrant expression or activity of the
 CC polypeptide, such as viral infection, in particular HIV or human T cell
 CC lymphotropic virus (HTLV) infection. The present sequence is a synthetic
 CC nucleotide binding sequence (zinc finger) expressed by a phage display
 CC library, which may be assembled into the zinc finger protein of the
 CC invention
 XX
 SQ Sequence 7 AA;
 Query Match 100.0%; Score 33; DB 6; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 QRANLRA 7
 Db 1 QRANLRA 7
 RESULT 5
 ABP53167
 ID ABP53167 standard; peptide; 8 AA.
 XX
 AC ABP53167;
 XX
 DT 12-NOV-2002 (first entry)
 XX
 DE Zinc finger nucleotide binding peptide SEQ ID NO:10.

XX Zinc finger binding domain; zinc finger nucleotide binding peptide;
 KW zinc finger; gene expression; modulation; promoter; viral; Lentivirus;
 KW human T-cell lymphocytic virus; HTLV; human immunodeficiency virus; HIV.
 XX
 OS Synthetic.
 XX
 XX WO200266640-A2.
 PN
 XX 29-AUG-2002.
 PD
 XX 21-FEB-2002; 2002WO-EP001862.
 PF
 XX 21-FEB-2001; 2001US-00791106.
 PR
 XX (NOVS) NOVARTIS AG.
 PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
 PA (SCRI) SCRIPPS RES INST.
 XX
 PI Barbas CF, Dreier B;
 XX WPI; 2002-674941/72.
 DR
 XX New polypeptides comprising 2 to 10 zinc finger-nucleotide binding
 PT peptides, useful for modulating gene expression in promoters from viral
 PT groups, e.g. Lentivirus group, such HTLV-1 and 2, or HIV-1 and 2.
 PT
 XX Claim 1; Fig 2; 48pp; English.
 PS
 XX The present invention describes a polypeptide comprising 2 to 12 zinc
 CC finger-nucleotide binding peptides in which at least one contains a
 CC nucleotide binding region. Also described: (1) an isolated and purified
 CC polynucleotide encoding the polypeptide cited above; (2) an expression
 CC vector containing the polynucleotide in (1); and (3) a process of
 CC regulating expression of a nucleotide sequence that contains the sequence
 CC SI comprising exposing of a nucleotide sequence to the polypeptide cited
 CC above. (SI) is (5'-ANN)n-3', where n = any integer from 2-12; and N = A,
 CC C, G, or T. The polypeptide is useful for modulating gene expression in
 CC promoters from viral groups, e.g. Lentivirus group, such as human T-cell
 CC lymphocytic virus (HTLV) 1 and 2, or human immunodeficiency virus (HIV) 1
 CC or 2. AB075687 to AB075697 and ABP53112 to ABP53221 represent sequences
 CC given in the exemplification of the present invention
 XX
 SQ Sequence 8 AA;
 Query Match 100.0%; Score 33; DB 5; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 QRANLRA 7
 Db 2 QRANLRA 8
 RESULT 6
 ABU60713
 ID ABU60713 standard; peptide; 8 AA.
 XX
 AC ABU60713;
 XX
 DT 06-MAY-2003 (first entry)
 XX
 DE Phage displayed zinc finger recognising AAA #4.
 XX
 KW Zinc finger; phage display; ant-HIV; virucide; HIV infection;
 KW human T cell lymphotropic virus infection; HTLV.
 XX
 OS Synthetic.
 XX
 XX US2002165356-A1.
 PN
 PD 07-NOV-2002.
 XX

PF 21-FEB-2002; 2002US-00080100.
XX
PR 21-FEB-2001; 2001US-00367356.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Barbas CF, Dreier B;
XX
XX WPI; 2003-255225/25.
DR
XX
XX New zinc finger-nucleotide binding polypeptides, useful in the treatment
PT of disorders associated with the aberrant expression or activity of the
PT polypeptide, such as HIV infection or human T cell lymphotropic virus
PT infection.
XX
XX Claim 1; Fig 2; 22pp; English.
PS
XX The invention relates a new polypeptide comprising from 2-12 zinc finger-
CC nucleotide binding peptides at least one of which contains a nucleotide
CC binding region having any of 71 nucleotide binding sequences, fully
CC defined in the specification. Also included are an isolated and purified
CC polynucleotide that encodes the polypeptide, an expression vector
CC containing the polynucleotide and a process of regulating expression of a
CC nucleotide sequence that contains the sequence (5'-ANN) n-3', where n is
CC an integer from 2-12, the process comprising exposing the nucleotide
CC sequence to the zinc finger protein of the invention. The methods and
CC compositions of the present invention, are useful for modulating zinc-
CC finger protein binding and hence regulating aspects of gene expression.
CC The zinc finger polypeptides are useful in the treatment of treatment of
CC disorders associated with the aberrant expression or activity of the
CC polypeptide, such as viral infection, in particular HIV or human T cell
CC lymphotropic virus (HTLV) infection. The present sequence is a synthetic
CC nucleotide binding sequence (zinc finger) expressed by a phage display
CC library, which may be assembled into the zinc finger protein of the
CC invention
XX
XX Sequence 8 AA;
SQ
Query Match 100.0%; Score 33; DB 6; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QRANLRA 7
Db 2 QRANLRA 8
|||||
RESULT 7
ADJ98368
ID ADJ98368 standard; peptide; 8 AA.
XX
XX AC ADJ98368;
XX
XX 06-MAY-2004 (first entry)
DT
XX Zinc finger DNA binding peptide #39.
DE
XX library; multimeric DNA binding polypeptide;
KW zinc finger DNA binding peptide; gene expression silencing;
KW gene expression enhancement.
XX
XX Unidentified.
OS
XX WO2003066828-A2.
XX
XX 14-AUG-2003.
PD
XX
XX 07-FEB-2003; 2003WO-US003705.
PF
XX
XX 07-FEB-2002; 2002US-0354981P.
PR
XX (SCRI) SCRIPPS RES INST.
PA
XX

PI Barbas CF, Blancafort P;
XX
XX WPI; 2003-731499/69.
XX
XX New zinc finger library of multimeric DNA binding polypeptides, useful
PT for sterically occluding the binding site of a natural transcription
PT factor, and enhancing or silencing target gene expression.
XX
XX Example 11; Fig 19; 64pp; English.
PS
XX The invention comprises a library of multimeric DNA binding polypeptides
CC (e.g. zinc finger DNA binding polypeptides). The zinc finger binding
CC polypeptides of the invention are useful for sterically occluding the
CC binding site of a natural transcription factor and enhancing or silencing
CC target gene expression. The present amino acid sequence represents a zinc
CC finger DNA binding peptide of the invention.
XX
XX Sequence 8 AA;
SQ
Query Match 100.0%; Score 33; DB 7; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QRANLRA 7
Db 2 QRANLRA 8
|||||
RESULT 8
ADK14951
ID ADK14951 standard; protein; 600 AA.
XX
XX AC ADK14951;
XX
XX 06-MAY-2004 (first entry)
DT
XX Urinary specific protein #57.
DE
XX cytostatic; vaccine; gene therapy; urinary specific nucleic acid;
KW diagnosis; urinary cancer; transgenic animal.
KW
XX Homo sapiens.
OS
XX WO2003057839-A2.
PN
XX 17-JUL-2003.
PD
XX 23-DEC-2002; 2002WO-US041613.
PF
XX 28-DEC-2001; 2001US-0343690P.
PR
XX (DIAD-) DIADEXUS INC.
PA
XX Macina RA, Sun Y, Liu C;
XX
XX WPI; 2003-587114/55.
DR
XX New urinary specific nucleic acid molecules and polypeptides, useful in
PT gene therapy, or for identifying, diagnosing, monitoring, staging,
PT imaging or treating urinary cancer and non-cancerous disease states of
PT the urinary tract.
XX
XX Claim 12; SEQ ID NO 330; 392pp; English.
PS
XX The invention relates to novel urinary specific genes and their encoded
CC proteins or a nucleic acid molecule that selectively hybridizes, or has
CC at least 95 % sequence identity, to these nucleic acid molecules. The
CC urinary specific nucleic acid molecules, polypeptides and antibodies
CC against the polypeptides are useful for identifying, diagnosing,
CC monitoring, staging, imaging or treating urinary cancer and non-cancerous
CC disease states of the urinary tract. The nucleic acids and polypeptides
CC are also useful in gene therapy, for designing and/or identifying
CC antagonists or agonists of the polypeptides and for producing transgenic

CC animals and cells. This sequence corresponds to one of the urinary
 CC specific proteins of the invention.
 XX
 XX Sequence 600 AA;

Query Match 100.0%; Score 33; DB 7; Length 600;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 7
 Db 444 QRANLRA 450
 |||||

RESULT 9

ID ABO83025 standard; protein; 177 AA.

AC ABO83025;

DT 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #15200.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

DR N-PSDB; ABD16596.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 31771; 455pp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a P. aeruginosa nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-P. aeruginosa drugs, as templates for recombinant production of P. aeruginosa-derived peptides or polypeptides, as target components for diagnosis and/or treatment of P. aeruginosa-caused infection, and in detection of P. aeruginosa sequences or other sequences of Pseudomonas species using biochip technology. Sequences ABO67826-ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html

XX Sequence 177 AA;

Query Match

Best Local Similarity 87.9%; Score 29; DB 7; Length 177;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 6
 Db 112 QRANLR 117
 |||||

RESULT 10

AAW70232
 ID AAW70232 standard; protein; 510 AA.

XX AAW70232;

DT 17-OCT-2003 (revised)

DT 13-NOV-1998 (first entry)

XX Leishmania antigen LcgSP3 protein.

XX Leishmania antigen; immune response; infection detection; therapy;

XX humoral response induction; cellular response induction; cancer;

XX interleukin-12 production.

XX Leishmania donovani chagasi.

XX WO9835045-A2.

PD 13-AUG-1998.

XX 12-FEB-1998; 98WO-US003002.

XX 12-FEB-1997; 97US-00798841.

PR 27-AUG-1997; 97US-00920609.

XX (CORI-) CORIXA CORP.

XX Reed SG, Campos-Neto A, Webb JR, Dillon DC, Skeiky YA;

XX WPI; 1998-447242/38.

DR N-PSDB; AAV47577.

XX New immunogenic fragments of Leishmania antigens and related nucleic acid, vectors and host cells - are useful for diagnosis, prevention and treatment of leishmaniasis, also to induce production of interleukin-12 generally.

PS Claim 13; Page 125-127; 194pp; English.

XX This sequence encodes a Leishmania antigen (LAG) of the invention, designated LcgSP3. Compositions and vaccines containing the protein are used to generate a protective or therapeutic immune response against the Leishmania species donovani, chagasi, infantum, major, amazonensis, braziliensis, panamensis, tropica or guayanensis. They can also be used to detect infection (in a skin test). The compositions induce a humoral and/or cellular response, specifically of Th1 type, particularly including induction of interleukin-12 (IL-12) production. They may thus be used more generally to treat any condition (e.g. bacterial, viral or protozoal infection, or cancer) which responds to IL-12. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 510 AA;

Query Match

Best Local Similarity 87.9%; Score 29; DB 2; Length 510;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 6

Db 114 QRANLR 119
 |||||

RESULT 11

AAE24946
 ID AAE24946 standard; protein; 510 AA.

XX AAE24946;

```

XX 29-AUG-2003 (revised)
DT 22-OCT-2002 (first entry)
XX
DE Leishmania chagasi LcgSP3 antigenic protein.
XX
XX Immunogen; Leishmania antigen; therapy; delayed-type hypersensitivity;
KW leishmaniasis; vaccine; interleukin-12 stimulation; cancer; protozoacide;
KW virucide; bactericide; cytostatic; immune response; LcgSP3 protein.
XX
XX Leishmania donovani chagasi.
XX
XX US6375955-B1.
XX
XX 23-APR-2002.
XX
XX 12-FEB-1998; 98US-00022765.
XX
XX 22-SEP-1995; 95US-00533669.
PR 12-FEB-1997; 97US-00798841.
XX 27-AUG-1997; 97US-00920609.
XX
XX (CORI-) CORIXA CORP.
XX
XX Read SG, Campos-Neto A, Webb JR, Dillon DC, Skeiky YAW, Bhatia A;
PI Probst P;
XX
XX WPI; 2002-424754/45.
DR N-PSDB; AD40304.
XX
XX New polypeptide comprising immunogenic part of Leishmania antigen, useful
PT for treatment, prevention and diagnosis of leishmaniasis.
XX
XX Example 14; Col 121-124; 111pp; English.
XX
XX The present invention relates to novel proteins comprising immunogenic
CC part of Leishmania antigen. Compositions containing sequences of the
CC invention and other Leishmania antigen related polypeptides are useful
CC for preventing, treating and detecting (in delayed-type hypersensitivity
CC skin tests) leishmaniasis. They can also be used to treat any diseases
CC responsive to interleukin-12 stimulation, including bacterial, viral and
CC protozoal infections and cancer. Sequences of the invention are useful as
CC vaccines. The present sequence is Leishmania chagasi LcgSP3 antigenic
CC protein. (Updated on 29-AUG-2003 to standardise OS field)
XX
XX Sequence 510 AA;
SQ
Query Match 87.9%; Score 29; DB 5; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLR 6
Db 114 QRANLR 119
|||||

RESULT 12
AAU71828
ID AAU71828 standard; protein; 510 AA.
XX
XX AAU71828;
XX
XX 29-AUG-2003 (revised)
DT 26-FEB-2002 (first entry)
XX
XX Leishmania antigen LcgSP3.
XX
XX Leishmaniasis; Leishmania antigen; immunostimulant; protozoacide; Ldp23;
KW interleukin-15; lbp83; M15; Lt-1; lbelF4A; LmsplA; LmgSP9A; MAPS-1A;
KW LmgSP1; LmgSP3; LmgSP5; LmgSP9; LmgSP13; LmgSP19; LcgSP1; LcgSP3;
KW LcgSP4; LcgSP8; LcgSP10; LG6-34; LE6-44; 4A5-63; 1B11-39; 2A10-37;
KW 4G2-83; 4H6-41; 8G3-100.
XX

OS Leishmania donovani chagasi.
PN WO200179276-A2.
XX
XX 25-OCT-2001.
XX
XX 05-APR-2001; 2001WO-US011254.
XX
XX 14-APR-2000; 2000US-00551974.
PR 05-MAY-2000; 2000US-00565501.
XX 14-AUG-2000; 2000US-00639206.
XX
XX (CORI-) CORIXA CORP.
XX
XX Read SG, Campos-Neto A, Webb JR, Dillon DC, Skeiky YAW, Bhatia A;
PI Coler RM, Probst P;
XX
XX WPI; 2002-061971/08.
DR N-PSDB; AAS96041.
XX
XX New isolated Leishmania antigens, useful for prevention, treatment and
PT diagnosis of leishmaniasis, also related nucleic acids for genetic
PT vaccination.
XX
XX Claim 1; Page 143-144; 193pp; English.
XX
XX The invention relates to polypeptides comprising an immunogenic part of a
CC Leishmania antigen. The Leishmania polypeptides and their associated DNA
CC sequences, epitopes and fusion proteins are used in the production of
CC compositions used for inducing a protective immune response against
CC leishmaniasis, for prevention and treatment of the disease. The
CC compositions can also be used generally to treat diseases that respond to
CC interleukin-15 stimulation. In addition, the products may contain an
CC immunostimulant. Sequences AAU71805-AAU71862 represent Leishmania
CC antigens and antigenic peptides of the invention. (Updated on 29-AUG-2003
CC to standardise OS field)
XX
XX Sequence 510 AA;
SQ
Query Match 87.9%; Score 29; DB 5; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRANLR 6
Db 114 QRANLR 119
|||||

RESULT 13
ABG60884
ID ABG60884 standard; protein; 510 AA.
XX
XX ABG60884;
XX
XX 29-AUG-2003 (revised)
DT 19-AUG-2002 (first entry)
XX
XX Leishmania antigenic polypeptide #21.
XX
XX Leishmania antigen; Th1; interleukin-12; IL-12; Leishmaniasis;
KW protozoacide; immune response enhancer; vaccine.
XX
XX Leishmania donovani chagasi.
XX
XX US6365165-B1.
XX
XX 02-APR-2002.
XX
XX 30-OCT-1998; 98US-00183861.
XX
XX 22-SEP-1995; 95US-00533669.
PR 12-FEB-1997; 97US-00798841.
XX 27-AUG-1997; 97US-00920609.
PR

```

PR 12-FEB-1998; 98US-00022765.
 XX (CORI-) CORIXA CORP.
 PA
 XX
 PI Reed SG, Campos-Neto A, Webb JR, Dillon DC, Skeiky YAW;
 XX WPI; 2002-433419/46.
 DR N-PSDB; ABK81750.
 XX
 PT Stimulating immune responses in a patient for preventing or treating
 PT Leishmaniasis, by administering a vaccine comprising a polypeptide
 PT comprising an immunogenic portion or epitope of Leishmania antigen.
 PS Claim 1; Col 123-126; 114pp; English.
 XX
 CC The invention relates to stimulating an immune response in a patient
 CC comprising administering to the patient a vaccine containing a non-
 CC specific immune response enhancer and a polypeptide having an immunogenic
 CC portion of Leishmania antigen, or a polypeptide having two contiguous
 CC epitopes of a Leishmania antigen. The method is useful for stimulating an
 CC immune response, in particular a Th1 response or interleukin-12 (IL-12)
 CC production in a patient, for preventing and treating Leishmaniasis.
 CC Sequences ABG60864-ABG60896 and ABG60903-ABG60910 represent Leishmania
 CC antigenic polypeptides and peptide fragments of the invention. (Updated
 CC on 29-AUG-2003 to standardise OS field)
 XX
 SQ Sequence 510 AA;
 Query Match 87.9%; Score 29; DB 5; Length 510;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ORANLR 6
 Db 114 ORANLR 119
 RESULT 14
 AAB71282
 ID AAB71282 standard; protein; 510 AA.
 AC AAB71282;
 XX
 DT 29-AUG-2003 (revised)
 DT 18-NOV-2002 (first entry)
 XX
 DE L. chagasi LcgSP3 antigen SEQ ID 50.
 XX
 KW Antigen; immunogenic; antiparasitic; immunostimulant; leishmaniasis;
 KW gene therapy; vaccine; interleukin-12 agonist.
 XX
 OS Leishmania donovani chagasi.
 XX
 XX
 FN US2002081320-A1.
 XX
 PD 27-JUN-2002.
 XX
 XX 04-JUN-2001; 2001US-00874923.
 XX
 PR 22-SEP-1995; 95US-00533669.
 PR 12-FEB-1997; 97US-00798841.
 PR 27-AUG-1997; 97US-00920609.
 PR 12-FEB-1998; 98US-00022765.
 PR 30-OCT-1998; 98US-00183861.
 PR 14-APR-2000; 2000US-00551974.
 PR 05-MAY-2000; 2000US-00565501.
 PR 14-AUG-2000; 2000US-00639206.
 XX
 XX (REED/) REED S G.
 PA (CAMP/) CAMPOS-NETO A.
 PA (WEBB/) WEBB J R.
 PA (DILL/) DILLON D C.
 PA (SKEI/) SKEIKY Y A W.

PA (BHAT/) BHATIA A.
 PA (COLE/) COLER R N.
 PA (PROB/) PROBST P.
 PA (BRAN/) BRANNON M.
 XX
 PI Reed SG, Campos-Neto A, Webb JR, Dillon DC, Skeiky YAW, Bhatia A;
 PI Coler RN, Probst P, Brannon M;
 XX WPI; 2002-635457/68.
 DR N-PSDB; AAF88543.
 XX
 PT New polypeptide containing at least an immunogenic portion of one or more
 PT Leishmania antigens or their variants, useful for preventing, treating
 PT and detecting leishmaniasis, and stimulating immune responses in
 PT patients.
 PS Claim 1; Page 60-61; 163pp; English.
 XX
 CC This invention describes a novel polypeptide containing an immunogenic
 CC portion of a Leishmania antigen or its variant which has antiparasitic
 CC and immunostimulant activity. The compositions and methods of the present
 CC invention are useful for preventing, treating and detecting
 CC leishmaniasis, and stimulating immune responses in patients against
 CC leishmaniasis. The polypeptides and the polynucleotides encoding them can
 CC be used for gene therapy, in vaccines or as interleukin-12 agonists. The
 CC compositions and methods of the present invention, as compared to prior
 CC art, are more improved therapeutic modalities in the diagnosis,
 CC prevention and treatment of leishmaniasis. (Updated on 29-AUG-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 510 AA;
 Query Match 87.9%; Score 29; DB 5; Length 510;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ORANLR 6
 Db 114 ORANLR 119
 RESULT 15
 ADB78813
 ID ADB78813 standard; protein; 510 AA.
 XX
 AC ADB78813;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Leishmania antigen LcgSP3.
 XX
 KW Antigen; protozoacide; antibacterial; virucide; cytostatic;
 KW immunostimulant; leishmaniasis; Leishmania infection; immune response;
 KW interleukin-2 stimulation; cancer; bacterial infection; viral infection;
 KW protozoan infection.
 XX
 OS Leishmania donovani chagasi.
 XX
 XX US2002169285-A1.
 PN 14-NOV-2002.
 PD
 XX 20-NOV-2001; 2001US-00991496.
 XX
 PR 22-SEP-1995; 95US-00533669.
 PR 12-FEB-1997; 97US-00798841.
 PR 27-AUG-1997; 97US-00920609.
 PR 12-FEB-1998; 98US-00022765.
 PR 30-OCT-1998; 98US-00183861.
 PR 14-APR-2000; 2000US-00551974.
 PR 05-MAY-2000; 2000US-00565501.
 PR 14-AUG-2000; 2000US-00639206.
 PR 04-JUN-2001; 2001US-00874923.

XX (REED/) REED S G.
PA (CAMP/) CAMPOS-NETO A.
PA (WEBB/) WEBB J R.
PA (DILL/) DILLON D C.
XX
PI Reed SG, Campos-Neto A, Webb JR, Dillon DC;
XX
XX WPI; 2003-605673/57.
DR N-PSDB; ADB78808.
XX
XX Novel isolated polypeptide useful for preventing or treating
PT leishmaniasis, comprises an immunogenic portion of a Leishmania antigen
PT or its variant.
XX
XX Claim 1; Page 61-62; 183pp; English.
XX
CC The invention relates to an isolated polypeptide comprising an
CC immunogenic portion of a Leishmania antigen or its. Also included are
CC antigenic epitopes, fusion proteins comprising an isolated polypeptide
CC fusion protein comprising at least two contiguous antigenic epitopes,
CC polynucleotides encoding the antigens or fusion proteins, a recombinant
CC expression vector comprising the polynucleotide, a host cell transformed
CC with the vector and a composition (pharmaceutical or immunogenic)
CC comprising the antigen or fusion protein and a physiologically acceptable
CC carrier. The compositions are useful for inducing protective immunity
CC against leishmaniasis in a patient. The fusion protein is useful for
CC detecting Leishmania infection in a patient by contacting dermal cells of
CC the patient with the composition and detecting an immune response on the
CC patient's skin, where the immune response is induration. The compositions
CC are useful for stimulating a cellular and/or humoral immune response in a
CC patient, or for treating a patient with a disease responsive to
CC interleukin (IL)-2 stimulation, where the disease is cancer or an
CC infection such as bacterial, viral or protozoan infection. The antigen is
CC useful for preventing or treating leishmaniasis. The present sequence
CC represents a Leishmania antigen (or fragment).
XX
SQ Sequence 510 AA;

Query Match 87.9%; Score 29; DB 7; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRANLR 6
Db 114 QRANLR 119

Search completed: April 21, 2005, 13:04:12
Job time : 76.1 secs